

INPUT SET: S19158.raw

This Raw Listing contains the General Information Section and up to the first 5 pages.

SEQUENCE LISTING

3 (1) General Information:

5 (i) APPLICANT: Inouye, Sumiko
6 Hsu, Mei-Yin
7 Eagle, Susan
8 Inouye, Masayo

10 (ii) TITLE OF INVENTION: Prokaryotic Reverse Transcriptase

12 (iii) NUMBER OF SEQUENCES: 45

14 (iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Weiser & Associates
(B) STREET: 230 South Fifteenth Street, Suite 500
(C) CITY: Philadelphia
(D) STATE: Pennsylvania
(E) COUNTRY: U.S.A.
(F) ZIP: 19102

22 (v) COMPUTER READABLE FORM:

23 (A) MEDIUM TYPE: Floppy disk
24 (B) COMPUTER: IBM PC compatible
25 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
26 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

27
28 (vi) CURRENT APPLICATION DATA:
29 (A) APPLICATION NUMBER: 08/808,031
30 (B) FILING DATE: 03-MAR-1997
31 (C) CLASSIFICATION: 435

32
33 (vii) PRIOR APPLICATION DATA:
34 (A) APPLICATION NUMBER: US 08/269,118
35 (B) FILING DATE: 30-JUN-1994

37 (viii) ATTORNEY/AGENT INFORMATION:
38 (A) NAME: Weiser, Gerard J.
39 (B) REGISTRATION NUMBER: 19,763
40 (C) REFERENCE/DOCKET NUMBER: 377.5888P

46

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/808,031DATE: 07/22/97
TIME: 10:54:15

INPUT SET: S19158.raw

47 (2) INFORMATION FOR SEQ ID NO:1:

48

49 (i) SEQUENCE CHARACTERISTICS:

50 (A) LENGTH: 2176 base pairs
51 (B) TYPE: nucleic acid
52 (C) STRANDEDNESS: double
53 (D) TOPOLOGY: linear

54

55

56 (ix) FEATURE:

57 (A) NAME/KEY: CDS
58 (B) LOCATION: 640..2094

59

60

61 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

62	TCATCCGCGC GGACACCCCC TCCTACGTGC CCCCGACGC GGAGAGCGGC GTGGAGACGG	60
63	TGTACCGCGT TTCCCTGGAT GGTACACCTGG TGGCGGTGGA GTGGGGCCCG CGCACGGGCT	120
64	CGCCCGTCA CCAGCGGCTC TGTTTCGACT CGGATGCCGA AGCCCCCGGA GCCTACTTCG	180
65	CGCGCCTCGA GAAGTTGGCG GCTGACGGCT ACATCGACGC GGCTCGGCA TTGGTCTAAA	240
66	CCCTTCAACC ACGGCTCGGC CGCCACGCGC GGCCGGCAGG ACAGGTGCCGA CGAACAGACG	300
67	ACGACGTGCG CTTCACGCGC GAGCAGCCGA GAGAGGTCCG GAGTGCATCA GCCTGAGCGC	360
68	CTCGAGCGGC GGAGCGGGGT TGCGCCGCTC CGGTTGGAAT GCAGGACACT CTCCGCAAGG	420
69	TAGCCTGTT TCAGCTCTCT CCCTCCTAGG CACTACGGCC AGGGTGGTA GCGGAGCAA	480
70	CGACGCCACC GCCGTTTACC CACCCCGGCC GTAGTGCCTA GGAGGGGAGA GCCGGTGAGG	540
71	CTACCGTGCC CCAGGTAAGA TGGTGGTGCT TTCCCGGCCT CCGTCGACTG CTCGCGCCAT	600
72	GTCCCGTCTT CCATCGCCGC GCCCGCCCAA GGTGCAGAC ATG ACC GCC AGG CTG	654
73	Asp Thr Ala Arg Leu	
74	1 5	
75	GAC CCG TTC GTC CCC GCA GCT TCG CCG CAG GCC GTG CCC ACG CCC GAG	702
76	Asp Pro Phe Val Pro Ala Ala Ser Pro Gln Ala Val Pro Thr Pro Glu	
77	10 15 20	
78	Leu Thr Ala Pro Ser Ser Asp Ala Ala Ala Lys Arg Glu Ala Arg Arg	750
79	25 30 • 35	
80	CTC ACC GCT CCG TCG TCA GAC GCG GCC GCG AAG CGT GAA GCC CGC CGG	
81	Leu Thr Ala Pro Ser Ser Asp Ala Ala Ala Lys Arg Glu Ala Arg Arg	
82	40 45 • 50	
83	CTC GCG CAC GAA GCG TTG CTC GTC CGC GCG AAG GCC ATC GAC GAA GCG	798
84	Leu Ala His Glu Ala Leu Leu Val Arg Ala Lys Ala Ile Asp Glu Ala	
85	40 45 • 50	
86	GGC GGC GCC GAC GAC TGG GTG CAG GCG CAG CTC GTC TCC AAG GGG CTC	846

**RAW SEQUENCE LISTING
PATENT APPLICATION US/08/808,031**

DATE: 07/22/97
TIME: 10:54:19

INPUT SET: S19158.raw

100	Gly Gly Ala Asp Asp Trp Val Gln Ala Gln Leu Val Ser Lys Gly Leu			
101	55	60	65	
102				
103	GCG GTC GAG GAC CTG GAC TTC TCC AGC GCC TCC GAG AAG GAC AAG AAG	894		
104	Ala Val Glu Asp Leu Asp Phe Ser Ser Ala Ser Glu Lys Asp Lys Lys			
105	70	75	80	85
106				
107	GCC TGG AAG GAG AAG AAG AAG GCC GAG GCC ACC GAG CGC CGC GCG CTG	942		
108	Ala Trp Lys Glu Lys Lys Ala Glu Ala Thr Glu Arg Arg Ala Leu			
109	90	95	100	
110				
111	AAG CGT CAG GCG CAC GAG GCG TGG AAG GCC ACG CAC GTG GGC CAC CTG	990		
112	Lys Arg Gln Ala His Glu Ala Trp Lys Ala Thr His Val Gly His Leu			
113	105	110	115	
114				
115	GGC GCG GGC GTG CAC TGG GCG GAG GAC CGC CTG GCC GAC GCG TTC GAC	1038		
116	Gly Ala Gly Val His Trp Ala Glu Asp Arg Leu Ala Asp Ala Phe Asp			
117	120	125	130	
118				
119	GTG CCC CAC CGC GAG GAG CGC GCG GCC CGG GCC AAC GGC CTG ACG GAG CTG	1086		
120	Val Pro His Arg Glu Glu Arg Ala Arg Ala Asn Gly Leu Thr Glu Leu			
121	135	140	145	
122				
123	GAC TCC GCG GAG GCG CTG GCC AAG GCG CTG GGG CTG AGC GTC TCC AAG	1134		
124	Asp Ser Ala Glu Ala Leu Ala Lys Ala Leu Gly Leu Ser Val Ser Lys			
125	150	155	160	165
126				
127	CTC CGC TGG TTC GCG TTC CAC CGG GAG GTC GAC ACG GCC ACG CAC TAC	1182		
128	Leu Arg Trp Phe Ala Phe His Arg Glu Val Asp Thr Ala Thr His Tyr			
129	170	175	180	
130				
131	GTG AGC TGG ACC ATT CCG AAG CGG GAC GGC AGC AAG CGC ACG ATT ACG	1230		
132	Val Ser Trp Thr Ile Pro Lys Arg Asp Gly Ser Lys Arg Thr Ile Thr			
133	185	190	195	
134				
135	TCC CCC AAG CCT GAG CTG AAG GCA GCG CAG CGC TGG GTG CTG TCC AAC	1278		
136	Ser Pro Lys Pro Glu Leu Lys Ala Ala Gln Arg Trp Val Leu Ser Asn			
137	200	205	210	
138				
139	GTC GTG GAG CGG CTG CCG GTC CAC GGC GCC GCC CAC GGC TTC GTG GCG	1326		
140	Val Val Glu Arg Leu Pro Val His Gly Ala Ala His Gly Phe Val Ala			
141	215	220	225	
142				
143	GGA CGC TCC ATC CTC ACC AAC GCG CTG GCC CAC CAG GGC GCG GAC GTC	1374		
144	Gly Arg Ser Ile Leu Thr Asn Ala Leu Ala His Gln Gly Ala Asp Val			
145	230	235	240	245
146				
147	GTG GTC AAG GTG GAC CTC AAG GAC TTC TTC CCC TCC GTC ACC TGG CGC	1422		
148	Val Val Lys Val Asp Leu Lys Asp Phe Phe Pro Ser Val Thr Trp Arg			
149	250	255	260	
150				
151	CGG GTG AAG GGC CTG TTG CGC AAG GGC GGC CTG CGG GAG GGC ACG TCC	1470		
152	Arg Val Lys Gly Leu Leu Arg Lys Gly Gly Leu Arg Glu Gly Thr Ser			

**RAW SEQUENCE LISTING
PATENT APPLICATION US/08/808,031**

DATE: 07/22/97
TIME: 10:54:22

INPUT SET: S19158.raw

153	265	270	275	
154				
155	ACG CTG CTG TCC CTC CTC TCC ACG GAA GCG CCG CGG GAG GCG GTC CAG			1518
156	Thr Leu Leu Ser Leu Leu Ser Thr Glu Ala Pro Arg Glu Ala Val Gln			
157	280	285	290	
158				
159	TTC CGC GGC AAG CTC CTG CAC GTC GCC AAG GGC CCG CGC GCC CTG CCC			1566
160	Phe Arg Gly Lys Leu Leu His Val Ala Lys Gly Pro Arg Ala Leu Pro			
161	295	300	305	
162				
163	CAG GGC GCC CCC ACG TCG CCC GGC ATC ACC AAC GCG CTC TGC CTG AAG			1614
164	Gln Gly Ala Pro Thr Ser Pro Gly Ile Thr Asn Ala Leu Cys Leu Lys			
165	310	315	320	325
166				
167	CTC GAC AAG CGG CTG TCC GCC CTC GCG AAG CGG CTG GGC TTC ACC TAC			1662
168	Leu Asp Lys Arg Leu Ser Ala Leu Ala Lys Arg Leu Gly Phe Thr Tyr			
169	330	335	340	
170				
171	ACG CGC TAC GCG GAC CTG ACC TTC TCC TGG ACG AAG GCG AAG CAG			1710
172	Thr Arg Tyr Ala Asp Asp Leu Thr Phe Ser Trp Thr Lys Ala Lys Gln			
173	345	350	355	
174				
175	CCC AAG CCG CGG CGG ACG CAG CGT CCC CCC GTC GCG GTC CTC CTG TCT			1758
176	Pro Lys Pro Arg Arg Thr Gln Arg Pro Pro Val Ala Val Leu Leu Ser			
177	360	365	370	
178				
179	CGC GTC CAG GAA GTG GTG GAG GCG GAG GGC TTC CGC GTG CAC CCG GAC			1806
180	Arg Val Gln Glu Val Val Glu Ala Glu Gly Phe Arg Val His Pro Asp			
181	375	380	385	
182				
183	AAG ACG CGC GTC GCC CGC AAG GGC ACG CGG CAG CGG GTC ACC GGG CTC			1854
184	Lys Thr Arg Val Ala Arg Lys Gly Thr Arg Gln Arg Val Thr Gly Leu			
185	390	395	400	405
186				
187	GTC GTG AAT GCG GCG AAG GAC GCG CCC GCG GCC CGA GTC CCG CGC			1902
188	Val Val Asn Ala Ala Gly Lys Asp Ala Pro Ala Ala Arg Val Pro Arg			
189	410	415	420	
190				
191	GAC GTC GTC CGC CAG CTC CGC GCC ATC CAC AAC CGG AAG AAG GGC			1950
192	Asp Val Val Arg Gln Leu Arg Ala Ala Ile His Asn Arg Lys Lys Gly			
193	425	430	435	
194				
195	AAG CCG GGC CGC GAG GGC GAG TCG CTC GAG CAG CTC AAG GGC ATG GCC			1998
196	Lys Pro Gly Arg Glu Gly Glu Ser Leu Glu Gln Leu Lys Gly Met Ala			
197	440	445	450	
198				
199	CCC TTC ATC CAC ATG ACG GAC CCG GCC AAG GGC CGC GCC TTC CTG GCT			2046
200	Ala Phe Ile His Met Thr Asp Pro Ala Lys Gly Arg Ala Phe Leu Ala			
201	455	460	465	
202				
203	CAG CTC ACG GAG CTC GAG TCC ACG GCG AGC GCC GCT CCG CAG GCG GAG			2094
204	Gln Leu Thr Glu Leu Glu Ser Thr Ala Ser Ala Ala Pro Gln Ala Glu			
205	470	475	480	485

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/808,031DATE: 07/22/97
TIME: 10:54:26

INPUT SET: S19158.raw

206 TGACGCTCAG CGCGCGTCCG TCGCCGACGT GCCGGCGGCC AGCAACGCCG CATTAGCAA 2154
207
208
209 CTCCGTCAGC CGGCGCGGGT AC 2176
210
211
212 (2) INFORMATION FOR SEQ ID NO:2:
213
214 (i) SEQUENCE CHARACTERISTICS:
215 (A) LENGTH: 263 amino acids
216 (B) TYPE: amino acid
217 (D) TOPOLOGY: linear
218
219 (ii) MOLECULE TYPE: protein
220
221
222
223 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
224
225 Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val Lys Gln Trp Pro
226 1 5 10 15
227 Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile Cys Thr Glu Met
228 20 25 30
229
230 Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu Asn Pro Tyr Asn
231 35 40 45
232
233 Thr Pro Val Phe Ala Ile Lys Lys Asp Ser Thr Lys Trp Arg Lys
234 50 55 60
235
236 Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu
237 65 70 75 80
238
239 Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys Lys Lys Ser
240 85 90 95
241
242 Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Asp
243 100 105 110
244
245 Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser Ile Asn Asn
246 115 120 125
247
248 Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp
249 130 135 140
250
251 Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr Lys Ile Leu Glu
252 145 150 155 160
253
254 Pro Phe Lys Lys Gln Asn Pro Asp Ile Val Ile Tyr Gln Tyr Met Asp
255 165 170 175
256
257 Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His Arg Thr Lys
258

PAGE: 1

**SEQUENCE VERIFICATION REPORT
PATENT APPLICATION US/08/808,031**

DATE: 07/22/97
TIME: 10:54:29

INPUT SET: S19158.raw

Line

Error

Original Text